Digital Innovation and Transformation: a quasi-systematic literature review

Sílvia Bogéa Gomes
University of Lisbon / INOV INESC Inovação
Lisbon, Portugal *silvia.bogea@inov.pt*

Miguel Mira da Silva
University of Lisbon / INOV INESC Inovação
Lisbon, Portugal *mms@tecnico.ulisboa.pt*

Flávia Maria Santoro
University of the State of Rio de Janeiro
Rio de Janeiro, Brasil *flavia@ime.uerj.br*

Abstract

Nowadays there are several examples of successful companies that run innovative digital business models. Studies indicate that companies that do not follow the technological tendencies will possibly cease to exist in the next years. Besides, Digital Transformation has a direct impact on relations and forms of consumption. However, although much have been said about this topic, the literature has not established yet a common ground about the meaning of Digital Innovation and Digital Transformation. The purpose of this paper is to identify, analyze and synthesize the various aspects of the main concepts related to Digital Innovation and Transformation (DI&T). We have done a quasi-systematic review of the literature, generating as a primary outcome a list of the main constructs related to DI&T, as well as their definitions. Our main contribution is a map that conceptualizes and relates DI and DT that could be used as a base for future researchers.

Keywords: Digital Innovation, Digital Transformation, Systematic Review, Concept.

1. Introduction

Uber[1], Airbnb[2], Facebook[3], LinkedIn[4], eBay[5] are some examples of successful and high-valued companies that have developed innovative digital business models. Both new companies arising from Digital Innovation and existing companies that have self-adapted to the digital age are growing fast into the market. According to Fichman *et al.* [5], after the popularization of the internet with the creation of a relatively cheap and increasingly easy-to-use world-wide digital infrastructure of computers, mobile devices, broadband network connections, and advanced application platforms, the digital technologies are being incorporated into our day-to-day life. They are introducing new habits, ways of living together, cultures, shapes for traditional businesses, and novel kinds of companies. Oswald and Kleinemeier [13] affirm that in the next ten years, 40% of companies indexed as Standard & Poor's 500 will cease to exist unless they follow the technological trends. Additionally, Oswald and Kleinemeier [13] observe that Digital Transformation has a straightforward impact on relations and forms of consumption.

This scenario has brought increasing interest in the Digital Innovation and/or Digital Transformation also as research fields. The attention to this subject can be observed, for

1 https://uberportugal.pt/portugal/
2 https://www.airbnb.com.br/
3 https://www.facebook.com/
4 https://www.linkedin.com
5 https://www.ebay.com

example, in a quick search in Google Scholar. Documents with "Digital Innovation" or "Digital Transformation" in the title or anywhere in the text, first until 2009, and then in the period from 2010 to2018, not including patents and citations, show a significant evolution in these two periods scenarios.

Up to 2009, approximately 4310 papers have been published, of which only 2% (approximately 85 papers) have these terms in their title. From 2010 to 2018, this amount of publications has been multiplied by at least seven times (to date, 28812 papers have been published approximately) and the number of papers that have these terms in the title has tripled, representing 6% of papers (approximately 1693 papers). Yet, during the last eight years, there has been a significant increase in the result of selected papers year-by-year as presented in Fig. 1.

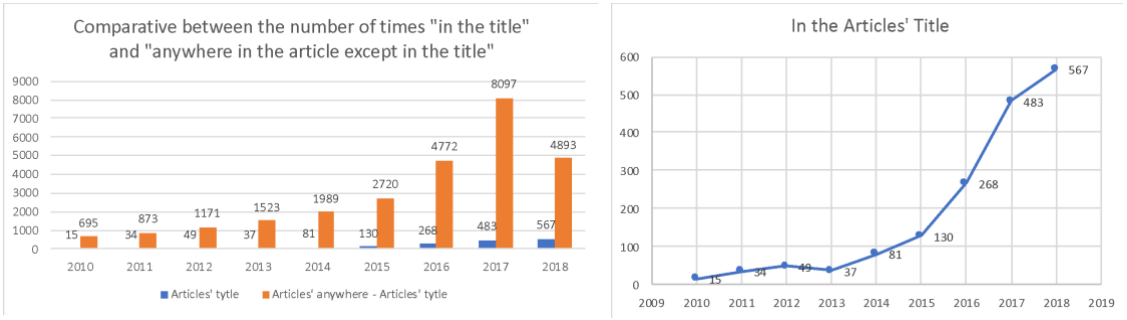

Fig. 1. The terms DI and DT: title x anywhere in the papers

However, there seems to be still no consensus about Digital Innovation (DI) or Digital Transformation (DT) concepts. For example, [2] observe that digitization, digitalization, and Digital Transformation are buzzwords in the changing context of society, working life and people behavior. The appropriate conceptualization of these terms is not yet consolidated, although existing publications discuss their benefits and consequences. Morakanyane *et al.* [12] state that even though the Digital Transformation has gained great research interest, the lack of common understanding of this concept is evident. Based on these examples, we argue that there is a need for a formalization of such concepts in order to establish a common ground for this area. Thus, we set out to contribute with this field with a *Quasi-Systematic* Literature Review (QSLR) about the main concepts regarding DI and DT. The context of our investigation was the academic and industrial publications.

According to Petersen *et al.* [14], this kind of study is also known as the Systematic Mapping Study, i.e. a study that provide an overview of the field's scope and enable discovering gaps and trends about the research field. So, we conducted the systematic mapping study, considering some practices of systematic review guidelines as well as those recommended by Petersen *et al.* [14]. The QSLR consists of four main phases: Review Design, Planning Review Process, Conducting Review Process and Reporting. With the aim of supporting the planning and to execution of the systematic mapping study, we used the online tool Parsif[6].

For Hinins *et al.* [8], DT is the combined effects of one or more DIs that by bringing novel actors, structures, practices, values, and beliefs that change, threaten, replace or complement existing rules within organizations, ecosystems, industries or field. We are following this definition, and as so, we decided to analyze the DT and DI main conceptual elements and their meanings together in this research. The goal of this paper is to identify, analyze and synthesize the diverse aspects of the main concepts related to DT and DI discussed in the literature. Our findings contribute to a better understanding of this research field in general. Thus, the research question investigated is: What are the constructs that contribute to conceptualize Digital Innovation and Digital Transformation?

The paper is organized as follows: Section 2 presents related work; Section 3 describes

6 https://parsif.al/

the methodology used for the QSLR; Section 4 presents the results obtained in detail; Section 5 discusses those results. Finally, Section 6 concludes the paper and explains limitations and suggestions for future research.

2. Related Work

We found fifteen systematic reviews about Digital Innovation and Digital Transformation. The twelve selected SLRs deal mainly with Conceptualization, Cases, Future and Challenges, as well as the identification of families of digital technologies that support the innovation process. Six SLR address Conceptualization, each of them focuses in some of concepts and the sum of the SLRs don't cover all. The goal of Reis *et al.* [16] was to propose a definition for Digital Transformation, delivering a general overview of the literature, along with some suggestions for future research. Morakanyane *et al.* [12] also studied the phenomenon of Digital Transformation, detailing what it is; how it behaves; what drives it; what impacts it creates, as well as where the impacts are felt. As the outcome, they developed a concept centric matrix and constructed a more inclusive and general definition. Bockshecker *et al.* [2] searched the conceptualization of the terms and the systematization of the phenomena connected to digitization, digitalization, and Digital Transformation. Bohnsack *et al.* [3] consolidated the state of art on Digital Transformation, by synthesizing existing research, revealing connections in the diversified literature, and identifying important gaps in our understanding. Their result was organized in three building blocks of Digital Transformation distinguished and delineated: Determinants, Processes and Outcomes. Tesch *et al.* [18] reviewed the literature aiming at the applicability of digital Business Model Innovation (BMI) projects, by elaborating tools and methodologies categorization, considering two major logics of evaluation: Analytical/effectual and quantitative/qualitative. Vesti *et al.* [19] defined disruption in the digital domain, also known as digital disruption.

From the two SLR about Cases, Vukšić *et al.* [21] examined the Digital Transformation case studies collection in the last decade from practice and explored how organizations perceive Digital Transformation, regarding time, country of case, type of industry and Digital Transformation focus. Cziesla [4] explored the diverse aspects of IT-enabled transformation in the financial service industry.

Among the four SLR that address Challenges and Future, Lund [10] investigated what are the challenges of actors' interactions in Digital Innovation and how they can be addressed. Henriette *et al.* [7] proposed a research agenda for Digital Transformation in a managerial perspective. Hausberg *et al.* [6] provided an overview of existing research on the DT from a business perspective. Finally, Mazza [11] identified which families of digital technologies are the basis for the innovation process, and which of them drive and are necessary for the implementation of the innovation activity in companies.

The consolidated result from the twelve systematic reviews is summarized in Table 1. None of the reviews found could answer our research question, since they do not present the main Digital Innovation and Digital Transformation conceptual elements and their meanings.

Table 1. The Summary of the systematic reviews identified and selected.

Title	DT	DI	Ref
A literature review on Digital Transformation in the financial service industry	1		[4]
A Preliminary Literature Review of Digital Transformation Case Studies	1		[21]
Activities to address challenges in Digital Innovation		1	[10]
Conceptualizing Digital Transformation in Business Organizations: A Systematic Review of Literature	1		[12]
Digital Transformation in Business Research: A Systematic Literature Review and Analysis	1		[6]
Digital Transformation: A Literature Review and Guidelines for Future Research	1		[16]
Same, same, but different!? A systematic review of the literature on Digital Transformation	1		[3]
Structured Literature Review of disruptive innovation theory within the digital domain		1	[19]
Systematization of the Term Digital Transformation and its Phenomena from a Socio-Technical Perspective--A Literature Review	1		[2]
The evaluation aspect of digital business model innovation: A literature review on tools and methodologies		1	[18]
The role of digital technology in the innovation process of companies: a systematic literature review in the innovation management field		1	[11]
The shape of Digital Transformation: a systematic literature review	1		[7]

3. Research method: the process of a *Quasi-Systematic* Literature Review

3.1. Review Design

Recker [15] states that "in all scientific research the concepts, constructs, and measurements should be as carefully and precisely defined as possible to allow others to use, apply, and challenge the definitions, concepts, and results in their own work". This motivated our goal to identify the main concepts that characterize DT and DI in order to have a consolidated basis that can be shared, understood and used. According to Recker [15], "A concept describes an abstract or general idea inferred or derived from specific instances that we perceive in the real world. Concepts are thus mental representations that we develop, typically based on Experience". Moreover, the author states that the concepts are abstract units of meaning and play a key role in the development and testing of scientific theories. Concepts could have many meanings. Concepts can be linked to one another via propositions. A proposition (conceptual hypothesis) need to be converted into an operational hypothesis to be tested. For Recker [15], "conclusively, there are many, potentially unlimited ways in which such a proposition could be tested – and in turn, many different results could be obtained". The abstract meaning of a concept plus its operationalization for something in the real world that can be measured, is termed construct.

3.2. Planning Review Process

The Planning Review Process Phase represents the structure of the study. At the beginning of this phase, we defined the study goals and research question. The findings should contribute to a better understanding of the DI and DT definitions. The goal of this QSLR leads to the corresponding research question: (RQ): What are the constructs that contribute to conceptualize DI and DT? To characterize the papers that we were looking for, we specified a search string that included the words "concept", "definition" and "construct", "strategy" (linked with general notion or idea), "model" and "ontology" and "new". We intended to collect papers that described successful cases (ones which added value). Moreover, we would like to know what literature have already learned about it, and because of this, we included the terms "systematic review" and "systematic literature review". In summary, the final search string for this QSLR was:

> Title:("digital transformation" OR "digital innovation") AND Abstract:((systematic AND review) AND (concept OR ontology OR definition OR model OR new OR value OR strategy OR success OR construct))

We searched the literature reviews, case/experience reports, papers involving concepts/definitions, business processes, conceptual models/ontologies on the subject, to have a panorama of what other researchers working in this area have already advanced. On the other hand, we were not interested in papers that do not have in their scope the concepts with their meanings and definitions, whether in a language other than English, have already been identified from another source (duplicated) or we could access the full text of the paper. The inclusion and exclusion criteria to select potentially relevant studies like for review were:

- **Inclusion Criteria**: Cases / Experiences; Focus on Business Process; Focus on concepts / definitions; Focus on conceptual models / ontologies; Literature review.
- **Exclusion Criteria**: Duplicated document; Full text not available; Not written in English; Out of scope; Referenced paper (only citation).

We focused on the following databases as the most relevant for the BPM and Information Systems (IS) discipline: ACM Digital Library, IEEE Digital Library, Scholar Google, Science@Direct, Scopus and Springer Link. We limited the search to the title, abstract and keywords, but we did not restrict the publication year. Besides journal and conference papers, we also included other documents such as thesis and technical reports.

3.3. Conducting Review Process

The search string was defined according to the rules of each source and executed them. 733 papers were found. In this next step, we identified all the 150 duplicated papers that came from different sources. Therefore, the abstracts needed be read to eliminate papers that did not deal with the subject at all, narrowing the number down to 76 pre-selected papers, as presented in Table 2. Then, we sorted the papers according to their similarities and applied the inclusion criteria to start analyzing the results.

Table 2. Number of papers selected by criteria.

Accepted		**Rejected**	
Literature review	13	Full text not available	19
Focus on concepts / definitions	18	Not written in English	1
Focus on conceptual models / ontologies	18	Out of scope	483
Cases / Experiences	10	Referenced paper (only citation)	4
Focus on Business Process	17	**Total**	**507**
Total	**76**		
Duplicated		**Total**	**150**

3.4. Reporting

We read the 76 documents looking for statements of conceptualizations, their meanings, and definitions related to Digital Innovation and Digital Transformation. Then, we identified 56 articles to work with.

With these 56 articles selected, we started a process to identify the constructs that contribute to conceptualize Digital Innovation and Digital Transformation. This process, presented in the Fig. 3, below, is based on some of the preprocessing techniques for text mining [20], as summarized in Figure 2. Vijayarani *et al.* [20] define text mining as the process of discovering useful information from text documents, also called knowledge discovery in text (KDT) or knowledge of intelligent text analysis.

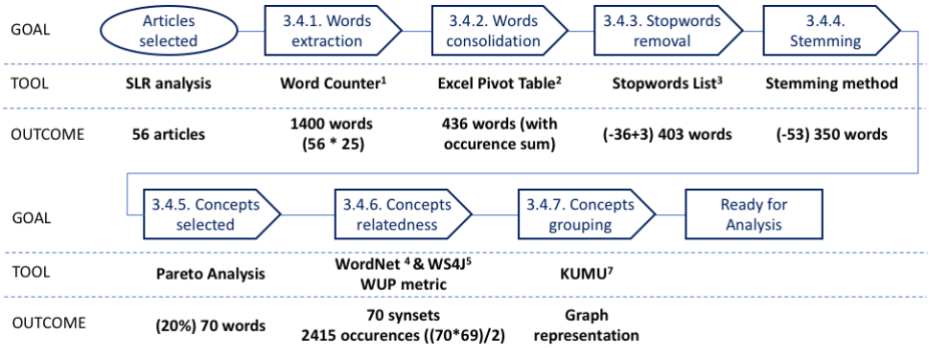

Fig. 2. The preprocessing for constructs analysis based on [19]

3.4.1. Words extraction

We selected a list the first 25 words that most appear in the texts (the minimum number of words offered), because the more words we select from each text (50, 100, ...), the bigger the set of words is highlighted, and the result could be dispersed. So, for each of the 56 papers, the Wordcounter presented the 25 most used words and how often each of the 25 words appeared in the paper. It resulted in a list of 1400 words (56 papers * 25 words) and the number of occurrences in each paper.

3.4.2. Words consolidation

As expected, some words appeared in more than one paper. So, we consolidated the 1400 words in a list of 436 different words that appeared in the 56 papers (using excel Pivot Table[7]), with the sum of the number of occurrences that each word appeared in all 56 papers.

3.4.3. Stopwords removal

Stop words (or stopwords) are the most common words in a language. There are some lists of stop words used by natural language processing (NLP) tools. Although the Wordcounter had excluded the "link words" before, it is not enough. Additionally, we had to remove the stopwords from the 436 different words list, using the Long Stopword List[8] (as for example, able, home, see). Also, we removed the words Digital, Innovation, and Transformation, because they are the universe of our research and mandatory in the title of the research query. As an outcome, we had 36 stopwords eliminated. However, we maintained three words classified in the list of stopwords that for our study should not be considered as stopwords: information, new and value.

3.4.4. Stemming

According to Vukšić *et al.* [20], the purpose of this method is to remove various suffixes, to reduce the number of words, to have accurately matching stems. It is used to identify the root/stem of a word. Words that do not have the same meaning should be kept separated. So, we analyzed the words set to identify roots or stems. We found root cases, such as regard / regarding / regards, result / results, and so on; and also, equivalent cases, such as, etal / et-al, R&D / research and others. Moreover, we found e-mails site addresses, and trademarks, like Amazon, Google, Apple, and ThyssenKrupp. Vukšić *et al.* [20] consider that one of the stemming purposes is reducing a word different grammatical forms like its noun, adjective, verb, and adverb. In this study, we transform all the words that are not a noun to its noun equivalent (e.g., manage to management, develop to development) to allow the comparison at the next step.

3.4.5. Concepts selected

In the next step, we used the Pareto diagram, a well-known column chart, which orders the words by occurrence frequencies, from highest to lowest. It supported the prioritization of the words by taking into account the Pareto principle (80% of the consequences come from of 20% of the causes). In our case, we considered 20% of words that occur more times in the 56 papers. As we can observe in Table 3, in the total, 350 words occur 36616 times. The 20% of words that occur most frequently (70 words), has the sum of occurrences (27135 occurrences) that represent 74% of occurrences. Moreover, 80% of the remaining 279 words represent 9381 occurrences (26%). So, Pareto's Principle fits well in this case.

Table 3. Words organized in the Pareto's Principle.

	Word	Sum of # occurrencess	% of occurences
20%	70	27135	74%
80%	280	9481	26%
Total	350	36616	

The 70 words list, with their total number of occurrences in the 56 papers, are in Table 4.

[7] "A pivot table is a table of statistics that summarizes the data of a more extensive table (such as from a database, spreadsheet, or business intelligence program). This summary might include sums, averages, or other statistics, which the pivot table groups together in a meaningful way" (source: https://en.wikipedia.org/wiki/Pivot_table).
[8] https://www.ranks.nl/stopwords

Table 4. 70 Words List classified in #occurrence descending order.

Word	# occur	Word	# occur	Word	# occur	Word	# occur
business	2203	study	459	ecosystem	212	open	128
technology	1452	design	416	architecture	197	culture	119
model	1424	case	399	component	195	experience	119
newness	1360	information	385	employee	190	impact	116
organization	1111	bank	377	market	189	evaluation	112
service	1035	firm	367	layer	180	offer	111
customer	1012	change	363	support	171	path	110
value	894	knowledge	344	field	170	information technology	107
company	814	resource	317	theory	162	manufacture	107
product	791	network	303	recombination	157	capability	106
process	784	platform	296	review	150	focus	106
management	766	literature	271	actor	148	project	106
digitization	661	mobility	261	physical	142	activity	102
user	530	artifact	253	cluster	133	channel	101
development	522	industry	241	practice	133	definition	101
strategy	504	work	240	provide	131	institutional	101
data	494	expert	221	social	131	mature	101
system	493			framework	128		

3.4.6. Concepts relatedness

Aiming to analyze the relationship between these 70 words, we used semantic measures. According to Adhikari *et al.* [1], nowadays semantic measures are widely used to assess the semantic relationship closeness between elements (informally, how a word A is related to a word B). Besides, the authors consider two types of semantic measures: similarity and relatedness. Semantic similarity "considers only taxonomical relationships for measuring the semantic strength between two concepts, e.g., rafting and water polo both are similar because both are water sports." While the semantic relatedness "considers taxonomic and nontaxonomic relations (e.g., meronymy, functionality, cause-effect, etc.) between concepts, e.g., food poison and stomach pain both are related. Food poison is the cause of stomach pain." Summing up, in general we can say that words' similarity means how close to be synonyms the two words are. In the other hand, words' relatedness characterizes a larger set of potential relationships between words, which means different types of relationships. For instance, antonyms have a higher relatedness (they essentially belong to the same semantic type) but low similarity (they are not very similar at all). In this work, we focus on the semantic relatedness between two concepts.

In order to measure the semantic relatedness between the concepts selected in the papers, we chose the Princeton's English WordNet 3.0 lexical database[9]. Wordnet is an online Knowledge Base thesaurus that groups words together based on their meanings. "Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations". Besides, we selected the WS4J (WordNet Similarity for Java)[10] for measuring semantic similarity/relatedness between the words identified using the Wordnet. From the semantic relatedness/similarity metrics available in WS4J, we selected the WUP relatedness metrics to analyze the relationship between the concepts, because it calculates the relatedness metric by considering the depths of the two synsets in the wordnet taxonomies, along wth the depth of the LCS (Least Common Subsumer)[11].

The LCS of two concepts A and B is "the most specific concept which is an ancestor of both A and B". For instance: a car is an automobile, and an automobile is a vehicle; a boat is a vehicle; vehicle is an object. In this case, "automobile" is the parent (and also ancestor) of "car", while "vehicle" is an ancestor of "car". "Vehicle" is also an ancestor of "boat". In this case, the LCS of "boat" and "car" is "vehicle", since it's the most specific concept which is an ancestor of both "boat" and "car". Note that while "object" is a common subsumer of both

9 https://wordnet.princeton.edu/
10 http://ws4jdemo.appspot.com
11 https://stackoverflow.com/questions/18629469/what-is-least-common-subsumer-and-how-to-compute-it

"boat" and "car", it is not the least, since there is still a child of "object" (in this case it's "vehicle") which is also a common subsumer of both "car" and "boat". "Automobile" is not the least common subsumer since it's not an ancestor of "boat". Each of the 70 words has their correspondent synset identified, using the common meaning of the words in the papers's text.

The next step was comparing each of the 70 selected synset in WordNet with the other 69 synsets, generating a list 2415 occurrences (not 4830, since the concept order does not matter in this case). This list of 2415 occurrences has as attributes the synset 1, the synset 2, and the Degree of Relatedness between these two concepts, using WUP metric.

3.4.7. Concepts grouping

In this step, we selected Kumu[12] software which is considered as "a powerful data visualization platform that helps organize complex information in interactive relationship maps" to analyze this network. We imported the data (the 205 connections of the 70 elements with the relatedness metric greater than 60%) into Kumu software.

Kumu has four chart layouts: System, Stakeholder, SNA (Social Network Analysis), and Custom. Each of them has a different goal. We identified the SNA type as the more adequate for our work, as "the template lets you identify key influencers, power structures, funding flows and more". The Graph generated by Kumu, considering the words and their connections with relatedness more than 0.6, is presented in Fig. 3.

Fig. 3. Graph Visualization considering connections with relatedness more than 0.6.

From now on, the analysis was done using the synset path, from the root until the word. For instance, the word Business, has the 10 types of synset. Because of the meaning of the word business used in the papers, we selected the synset #2, that means:

- S: (n) commercial enterprise, business enterprise, business (the activity of providing goods and services involving financial and commercial and industrial aspects) "computers are now widely used in business"

And the synset path are:

[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < psychological_feature#n#1 < event#n#1 < group_action#n#1 < transaction#n#1 < commerce#n#1 < business#n#2

[2] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < psychological_feature#n#1 < event#n#1 < act#n#2 < group_action#n#1 < transaction#n#1 < commerce#n#1 < business#n#2

Based on this rational, we identified that 84% of the words are abstract entities ("the general concept formed by extracting common features from specific examples") and 16% of the words are physical entities ("an entity that has physical existence"), as presented in Fig. 4.

12 https://kumu.io/

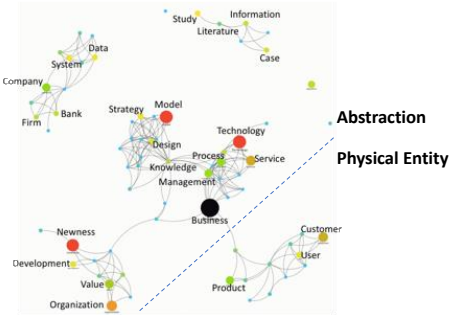

Fig. 4. Graph Visualization considering a high level of synset classification.

Following the same rational, we consolidated the words by their synset path (see Fig. 5), but detailing more, we could map the graph in two levels, the Cluster level (Fig. 6) and Construct level (Fig. 7). We are using Construct meaning as "concept, conception, construct (an abstract or general idea inferred or derived from specific instances)" and Cluster meaning as "a group of a number of similar things". Both definitions were derived from WordNet.

newness#n#1	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < property#n#2 < age#n#1 < newness#n#1
computer_architec	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < property#n#2 < constitution#n#4 < structure#n#2 < computer_architecture#n#2
resource#n#2	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < quality#n#1 < asset#n#1 < resource#n#2
focus#n#3	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < quality#n#1 < comprehensibility#n#1 < clarity#n#1 < focus#n#3
mobility#n#1	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < quality#n#1 < mobility#n#1
organization#n#5	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < quality#n#1 < regularity#n#2 < orderliness#n#1 < organization#n#5
value#n#2	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < quality#n#1 < worth#n#2 < value#n#2
activity#n#2	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < state#n#2 < activity#n#2
development#n#7	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < state#n#2 < condition#n#1 < improvement#n#3 < development#n#7
maturity#n#2	[1] *ROOT*#n#1 < entity#n#1 < abstraction#n#6 < attribute#n#2 < state#n#2 < maturity#n#2

Fig. 5. Synsets and their path.

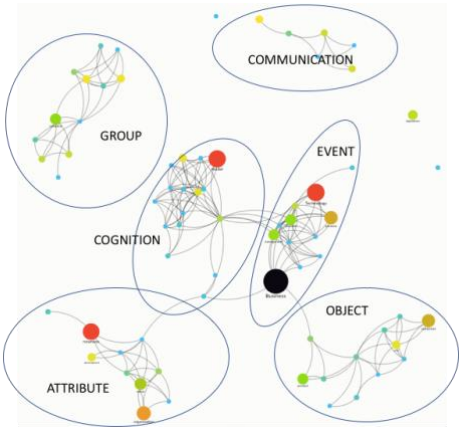

Fig. 6. Cluster Level of Graph Visualization.

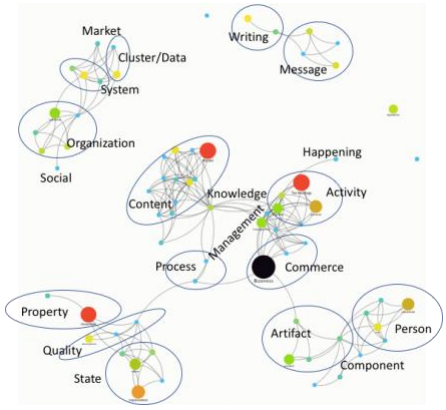

Fig. 7. Construct level of Graph Visualization.

In our study, we identified Group, Communication, Cognition, Attribute, Event and Object as Clusters. The Group and Communication are independent Clusters, but Cognition (result of learning and reasoning) and its Attributes are associated with the Object through the Event. The Constructs identified from the Clusters are detailed in Table 5.

Table 5. Clusters and Constructs identified.

CLUSTER	Construct	Definition (in WordNet)	Main characteristics (from 70 Words associated)
ATTRIBUTE		S: (n) attribute (an abstraction belonging to or characteristic of an entity)	
	Property	S: (n) property (a basic or essential attribute shared by all members of a class) "a study of the physical properties of atomic papers"	newness; computer_architecture
	Quality	S: (n) quality (an essential and distinguishing attribute of something or someone) "the quality of mercy is not strained"--Shakespeare	resource; focus; mobility; organization (as system synonym); value
	State	S: (n) state (the way something is with respect to its main attributes) "the current state of knowledge"; "his state of health"; "in a weak financial state"	activity; development; maturity
COMMUNICATION		S: (n) communication (something that is communicated by or to or between people or groups)	
	Message	S: (n) message, content, subject matter, substance (what a communication that is about something is about)	information; offer; definition; case
	Writing	S: (n) writing, written material, piece of writing (the work of a writer; anything expressed in letters of the alphabet (especially when considered from the point of view of style and effect)) "the writing in her novels is excellent"; "that editorial was a fine piece of writing"	study; literature
GROUP		S: (n) group, grouping (any number of entities (members) considered as a unit)	
	Collection	S: (n) collection, aggregation, accumulation, assemblage (several things grouped together or considered as a whole)	cluster, data
	Market	S: (n) market (the customers for a particular product or service) "before they publish any book they try to determine the size of the market for it"	
	Social	S: (adj) social, societal (relating to human society and its members) "social institutions"; "societal evolution"; "societal forces"; "social legislation"	
	Organization	S: (n) organization, organisation (a group of people who work together)	firm; industry; institution; company; bank
	System	S: (n) system, scheme (a group of independent but interrelated elements comprising a unified whole) "a vast system of production and distribution and consumption keep the country going"	system; ecosystem; network
COGNITION		S: (n) cognition, knowledge, noesis (the psychological result of perception and learning and reasoning)	
	Capability	S: (n) capability, capableness, potentiality (an aptitude that may be developed)	
	Culture	S: (n) culture (the attitudes and behavior that are characteristic of a particular social group or organization) "the developing drug culture"; "the reason that the agency is doomed to inaction has something to do with the FBI culture"	
	Support	S: (n) support (something providing immaterial assistance to a person or cause or interest) "the policy found little public support"; "his faith was all the support he needed"; "the team enjoyed the support of their fans"	
	Content	S: (n) content, cognitive content, mental object (the sum or range of what has been perceived, discovered, or learned)	theory; experience; model; framework; design; strategy; information_technology ; field
	Layer	S: (n) level, layer, stratum (an abstract place usually conceived as having depth) "a good actor communicates on several levels"; "a simile has at least two layers of meaning"; "the mind functions on many strata simultaneously"	
	Process	S: (n) process, cognitive process, mental process, operation, cognitive operation ((psychology) the performance of some composite cognitive activity; an operation that affects mental contents) "the process of thinking"; "the cognitive operation of remembering"	evaluation; review
	Open	S: (n) open, surface (information that has become public) "all the reports were out in the open"; "the facts had been brought to the surface"	
EVENT		S: (n) event (something that happens at a given place and time)	
	Activity	S: (n) activity (any specific behavior) "they avoided all recreational activity"	practice; process (synonym of

CLUSTER	Construct	Definition (in WordNet)	Main characteristics (from 70 Words associated)
			procedure); provision; technology; project; service
	Management	S: (n) management, direction (the act of managing something) "he was given overall management of the program"; "is the direction of the economy a function of government?"	
	Commerce	S: (n) commerce, commercialism, mercantilism (transactions (sales and purchases) having the objective of supplying commodities (goods and services)	business; manufacture; channel
	Happening	S: (v) happen, hap, go on, pass off, occur, pass, fall out, come about, take place (come to pass) "What is happening?"; "The meeting took place off without an incidence"; "Nothing occurred that seemed important"	change; recombination
OBJECT		S: (n) object, physical object (a tangible and visible entity; an entity that can cast a shadow) "it was full of rackets, balls and other objects"	
	Person	S: (n) person, individual, someone, somebody, mortal, soul (a human being) "there was too much for one person to do"	actor; expert; user; customer; employee
	Path	S: (n) path, track, course (a line or route along which something travels or moves) "the hurricane demolished houses in its path"; "the track of an animal"; "the course of the river"	
	Component	S: (n) component, constituent, element (an artifact that is one of the individual parts of which a composite entity is made up; especially a part that can be separated from or attached to a system) "spare components for cars"; "a component or constituent element of a system"	
	Artifact	S: (n) artifact, artefact (a man-made object taken as a whole)	artifact; product; work; platform

4. Discussion of Results

In Section 3.2, Planning Review Process, we defined the research scope as the Digital Context. Thus, we identified the papers whose title was concerned with “Digital Innovation" or "Digital Transformation". Based on those papers’ analysis, we identified the Clusters and Constructs from the relatedness words. In this section, we crosscheck our findings with the literature, and we observe how far the Clusters and Constructs identified are aligned.

Analyzing the Clusters and Constructs set identified from the papers selected, we could observe that **Digital Innovation & Transformation** is an **EVENT,** i.e., **Happening**, related to **Commerce**, **Management** and other **Activities**. Its **OBJECT** is a **Path** taken by **Persons** who act in various roles, generating **Artifacts** and its **Components**. Its **ATTRIBUTES** are characterized by newness and computer_architecture **Properties**, oriented to **Qualities** such as mobility and value, and **State** of Development, Maturity and Activities. The **COGNITION** is a fundamental element in this journey, it is **Open**, covers several **Layers**, and it is intensive in **Knowledge, Content** and **Process**, that **Supports** the new **Capabilities** development or the adequacy of existing ones, and also the **Culture** change. DI&T should be evaluated as a **GROUP**, a unit, analyzed as a **Collection** of clusters and data, in which the **Market**, the **Social** aspect, the **Organization** and the **System** (or the ecosystem) are central parts. The **COMMUNICATION** through messages/texts take place between the **GROUP**s and **OBJECT**s.

The literature agrees with that direction, as for Yoo *et al.* [22], Innovation is related to changes in thinking, products, processes or organizations. It is defined as the new elements’ introduction or an old elements’ new combination generating useful novelty (COGNITION and GROUP). It may also refer to incremental or radical changes. Sandhu [17] present Digital Revolution, Digital Business Transformation and Business Digital Transformation like Digital Transformation synonyms. For [12], Digital Transformation is “an organizational strategy formulated and executed by leveraging digital resources to create differential value” and they complement affirming that Digital Transformation is the integration of digital technologies into business processes (ATTRIBUTE, COGNITION and GROUP). Bohnsack *et al.* [3] concluded that Digital Transformation is a process organized into three building blocks - Determinants, Process and Outcome. For them, each building block has its own organization

(OBJECT, COGNITION, GROUP and COMMUNICATION). Ismail *et al.* [9] described the different Digital Transformation perspectives found in the literature. They represent them in concentric circles, where the outermost circle is related to the Digital Era perspective, followed respectively by the Social / Economic, Industry / Ecosystem, Network, Company / Institutional perspectives and finally, the innermost circle is related to the Individual (EVENT, ATTRIBUTE and GROUP).

Concerning the relationship between the terms Digital Transformation and Digital Innovation, Hinings *et al.* [8] observe that "Digital Transformation comes from the combined effects of several Digital Innovations bringing - for organizations, ecosystems, industries or fields - novel actors, structures, practices, values, and beliefs that change, threaten, replace or complement to existing rules (GROUP, OBJECT, COMMUNICATION and ATTRIBUTE).

5. Conclusion

The goal of this paper was to identify the main constructs related to Digital Innovation and Digital Transformation. The result obtained is a list of concepts that helps to provide understanding about DI&T. The DI&T subject is developing very quickly. So, a limitation of this research is that, when the review is redone, the results could be very different. But although the work could have expiration date, the process could be replicated.

From the list generated, we can explore papers related to each of the terms and enrich the definitions made so far. Likewise, we could use this Clusters and Constructs list as an initial guideline for DI&T initiatives, i.e., check whether the clusters and constructs are being explicitly considered. This may contribute to the completeness of your analysis. Thus, future research evolves the usage of this outcome produced. In the next step, we intend to develop a taxonomy based on those concepts, and moreover an ontology. Furthermore, we intend to evaluate the various adoptions of the DI&T classification, to show the diversity of existing angles in the current literature and support identifying future fields of investigation.

Acknowledgements

This work has been partially supported by Portuguese National funds through FITEC - Programa Interface, with reference CIT "INOV – INESC Inovação - Financiamento Base".

References

1. Adhikari, A., Singh, S., Dutta, A., Dutta, B.: A novel information theoretic approach for finding semantic similarity in WordNet. In: TENCON 2015-IEEE Region 10 Conference. pp. 1–6. IEEE (2015)
2. Bockshecker, A., Hackstein, S., Baumöl, U.: Systematization of the Term Digital Transformation and Its Phenomena From a Socio-Technical Perspective – a Literature Review. In: 26th European Conference on Information Systems (ECIS). , Portsmouth, UK (2018)
3. Bohnsack, R., Hanelt, A., Marz, D., Marante, C.: Same, same, but different!? A systematic review of the literature on digital transformation. In: Academy of Management Proceedings. p. 16262. , Briarcliff Manor, NY 10510 (2018)
4. Cziesla, T.: A literature review on digital transformation in the financial service industry. In: 27th Bled eConference: eEcosystems. pp. 25–36. , Bled, Slovenia (2014)
5. Fichman, R.G., Dos Santos, B.L., Zheng, Z. (Eric): Digital Innovation as a Fundamental and Powerful Concept in the Information Systems Curriculum. MIS Q. 38 (2), 329–353 (2014)
6. Hausberg, P., Liere-Netheler, K., Packmohr, S., Pakura, S., Vogelsang, K.: Digital Transformation in Business Research: A systematic literature review and analysis. In: DRUID18. , Copenhagen, Denmark (2018)
7. Henriette, E., Feki, M., Boughzala, I.: The Shape of Digital Transformation: A Systematic Literature Review. In: MCIS 2015 Proceedings. pp. 431–443. (2015)
8. Hinings, B., Gegenhuber, T., Greenwood, R.: Digital innovation and transformation: An

institutional perspective. Inf. Organ. 28 (1), 52–61 (2018)
9. Ismail, M.H., Khater, M., Zaki, M.: Digital Business Transformation and Strategy: What Do We Know So Far? Cambridge Serv. Alliance. November (2017)
10. Lund, J.: Activities to Address Challenges in Digital Innovation. In: Working Conference on Information Systems and Organizations. pp. 115–131. Springer Berlin Heidelberg (2014)
11. Mazza, E.M.: The role of digital technology in the innovation process of companies: a systematic literature review in the innovation management field. Politecnico Milano (2017)
12. Morakanyane, R., Grace, A., O'Reilly, P.: Conceptualizing Digital Transformation in Business Organizations: A Systematic Review of Literature. In: Bled eConference. p. 21. (2017)
13. Oswald, G., Kleinemeier, M.: Shaping the Digital Enterprise. Springer International Publishing, Cham (2017)
14. Petersen, K., Vakkalanka, S., Kuzniarz, L.: Guidelines for conducting systematic mapping studies in software engineering: An update. Inf. Softw. Technol. 64 1–18 (2015)
15. Recker, J.: Scientific Research in Information Systems. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
16. Reis, J., Amorim, M., Mel, N., Matos, P.: Digital Transformation: A Literature Review and Guidelines for Future Research. In: World Conference on Information Systems and Technologies. pp. 411–421. Springer, Cham (2018)
17. Sandhu, G.: The Role of Academic Libraries in the Digital Transformation of the Universities. In: 2018 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS). pp. 292–296. IEEE (2018)
18. Tesch, J., Brillinger, A., Tesch, J.F.: The Evaluation Aspect of Digital Business Model Innovation: A Literature Review on Tools and Methodologies. Proceedings of the 25th European Conference on Information Systems (ECIS) (2017)
19. Vesti, H., Nielson, C., Rosenstand, C.A.F., Massaro, M., Lund, M.: Structured Literature Review of disruptive innovation theory within the digital domain. In: ISPIM Innovation Symposium. The International Society for Professional Innovation Management (ISPIM). pp. 1–19. , Melbourne, Australia (2017)
20. Vijayarani, S., Ilamathi, J., Nithya: Preprocessing Techniques for Text Mining - An Overview. Int. J. Comput. Sci. Commun. Networks. 5 (1), 7–16 (2018)
21. Vukšić, V.B., Ivančić, L., Suša Vugec, D.: A Preliminary Literature Review of Digital Transformation Case Studies. In: ICMIT 2018: 20th International Conference on Managing Information Technology. pp. 737–742. (2018)
22. Yoo, Y., Henfridsson, O., Lyytinen, K.: The new organizing logic of digital innovation: An agenda for information systems research. Information Systems Research 21 (4), 724–735 (2010)